

## **General Disclaimer**

### **One or more of the Following Statements may affect this Document**

- This document has been reproduced from the best copy furnished by the organizational source. It is being released in the interest of making available as much information as possible.
- This document may contain data, which exceeds the sheet parameters. It was furnished in this condition by the organizational source and is the best copy available.
- This document may contain tone-on-tone or color graphs, charts and/or pictures, which have been reproduced in black and white.
- This document is paginated as submitted by the original source.
- Portions of this document are not fully legible due to the historical nature of some of the material. However, it is the best reproduction available from the original submission.

# Plausible Inference: A Multi-Valued Logic for Problem Solving

Leonard Friedman

(NASA-CR-158424) PLAUSIBLE INFERENCE: A N79-20750  
MULTI-VALUED LOGIC FOR PROBLEM SOLVING (Jet  
Propulsion Lab.) 26 p HC A03/MF A01  
CSCL 12A Unclass  
G3/59 17261

March 1, 1979

National Aeronautics and  
Space Administration

Jet Propulsion Laboratory  
California Institute of Technology  
Pasadena, California



# **Plausible Inference: A Multi-Valued Logic for Problem Solving**

**Leonard Friedman**

March 1, 1979

National Aeronautics and  
Space Administration

**Jet Propulsion Laboratory**  
California Institute of Technology  
Pasadena, California

**Page intentionally left blank**

## CONTENTS

1	A MULTI-VALUED LOGIC -----	2
2	SIMPLE CREDIBILITY PROPAGATION -----	4
2.1	QUANTITATIVE PLAUSIBLE INFERENCE -----	7
3	COMBINING CREDIBILITIES -----	9
3.1	DEFINING CRITERIA -----	10
3.2	COMBINING FUNCTIONS -----	11
4	REVISION OF OPINIONS -----	14
4.1	QUANTITATIVE REVISION RULES -----	14
4.2	A LEGAL EXAMPLE OF OPINION REVISION -----	16
4.2.1	Solution to the Legal Problem -----	16
5	SUCCESSIVE VERIFICATION -----	17
5.1	CREDIBILITY GROWTH THROUGH EXPERIENCE -----	18
5.2	A ROBOT LEARNING EXAMPLE -----	19
6	A MODEL OF HUMAN REASONING -----	20
6.1	FUZZY INTERFACES -----	20
6.2	PI IN PROBLEM SOLVING -----	21
	REFERENCES -----	22

### Table

1	SYNTAX OF BELIEF CHANGE -----	6
---	-------------------------------	---

"A person has a background, a machine has not. Indeed, you can build a machine to draw demonstrative conclusions for you, but I think you can never build a machine that will draw plausible inferences."

George Polya, in Patterns of Plausible Inference (Princeton University Press, Princeton, New Jersey, 1954), p. 116.

#### ABSTRACT

A new logic is developed which permits continuously variable strength of belief in the truth of assertions. Four inference rules result, instead of the two of formal logic, with formal logic as a limiting case. Quantification of belief is defined using the methods introduced by Shortliffe and Buchanan. Propagation of belief to linked assertions results from dependency-based techniques of truth maintenance so that local consistency is achieved or contradiction discovered in problem solving. Rules for combining, confirming, or disconfirming beliefs are given, and several heuristics are suggested that apply to revising already formed beliefs in the light of new evidence. The strength of belief that results in such revisions based on conflicting evidence appears to be a highly subjective phenomenon. Nevertheless, certain quantification rules appear to reflect an orderliness in the subjectivity. Several examples of reasoning by Plausible Inference (PI) are given, including a legal example and one from robot learning. Propagation of belief takes place in directions forbidden in formal logic and this results in conclusions becoming possible for a given set of assertions that are not reachable by formal logic.

## 1. A Multi-Valued Logic

One path to understanding the nature of intelligence is to compare various modes of reasoning. A kind of logic, which we shall call common-sense reasoning, is in constant use by all of us. It may produce guesses as output, and belief in assertions is expressed by a continuum of likelihoods rather than certainties. Formal logic, used by humans in special situations, proceeds instead with beliefs that are either certain or unknown. There have been many attempts by logicians, philosophers, and AI researchers to define new multi-valued logics in order to model aspects of common-sense reasoning with limited certainty. (See Shortliffe and Buchanan '75 for an introduction to the extensive literature produced by logicians and philosophers. In addition to Shortliffe '75, see Duda, Hart, Nilsson, and Sutherland '77 for several AI approaches.) My own work building a robot system for knowledge acquisition and learning from experience made me acutely aware of the need for a multi-valued logic. The robot learning system generates assertions whose truth is not certain and for which no formal probability of being true is known. It seemed highly desirable to find a logical system that could assign appropriate likelihoods of belief in new assertions and simultaneously maintain consistency of beliefs with related knowledge already stored.

This paper describes such a multi-valued logic arrived at by combining and generalizing the work of several investigators (Polya '54, Shortliffe and Buchanan '75, Thompson '79). Its form was suggested by the work of George Polya (Polya '54). His term for this type of reasoning was "plausible inference" and we shall apply it to the logic described here. Polya characterized assertions by qualitative belief values such as "more-likely" or "less-likely." These belief values could be most simply rendered quantitative by employing the methods introduced by Shortliffe and Buchanan (Shortliffe and Buchanan '75)— hereafter referred to as S&B. They gave simple rules for combining quantitative measures of belief in a hypothesis

supported by evidence. In their technique, as a starting point, a human expert supplied a numerical value for the belief in a hypothesis supported by a piece of confirming evidence. Note that in PROSPECTOR, investigators at the Stanford Research Institute (Duda, et al. '78) employ a somewhat different approach.

Still another line of research that provides an essential element to the logic of Plausible Inference has been the development of truth-maintenance in knowledge bases (Fikes '75, Stallman and Sussman '76, Doyle '78, London '79). Truth-maintenance systems automatically maintain the line of logical support for the truth or falsity of logically related assertions in a data base when a change in the truth of an assertion is introduced. Plausible Inference has been formulated so as to be compatible with the truth-maintenance system being built in our laboratory (Thompson '79). Thompson's approach has been to attach procedural specialists to the logical connectives (AND, OR, NOT, IMPLIES). In this formal logic system, each time the truth value of an assertion changes, logically connected assertions are examined by the appropriate specialist. This causes truth values to propagate through the knowledge base to maintain consistency, or discover either contradiction or lost support in a manner similar to other truth maintenance systems. Our implementation of plausible inference will proceed by augmenting this truth-maintenance system with additional belief values and modified procedural specialists.

We shall introduce several basic concepts and new notation, then review the four inference rules that are our starting point. These are Modus Ponens, Modus Tollens, the Method of Confirmation, and the Method of Weakened Support. This leads to a definition of "credibility propagation" for the four possibilities of antecedent truth-or-falsity, or consequent truth-or-falsity in a single implication. Rules for adding together confirming and disconfirming credibilities occurring at the same time can be given, but new phenomena seem to take over



when "opinions" have been formed and must be revised in the light of new evidence. (In our model, an opinion formed is equivalent to a non-zero credibility assignment.) When considering populations of linked assertions that require revision of a previously formed belief, we discover that additional combining rules are needed to model the propagation of credibility. No general theory for computing credibility revisions is known to us, and we cannot do more here than examine a few of the possibilities for this case. It is an area in which the discipline of psychology and our own introspection should offer many clues. We shall give an example of legal reasoning based on some of the heuristics we propose. A method for gradually increasing credibility assignment from a "don't-know" level to an appropriate probabilistic measure will be introduced for a special case of learning. The robot, during learning sessions, must constantly be revising already formed opinions, and modelling these situations may lead to additional insights into the phenomenon of how opinions change over the course of time.

## 2. Simple Credibility Propagation

Each assertion in the knowledge base has an associated property called "Truth Value" (TV), with a value set of true, false, or unknown {T, F, U}. In Plausible Inference, we introduce an additional property called "Credibility" which represents the degree-of-confidence in the truth or falsity of the assertion. Either truth value or credibility will be referred to as a belief. Let  $C(A)$  be the credibility of any assertion A. The value of  $C(A)$  is a number that ranges from -1 to +1 with 0 representing the unknown, +1 representing complete certainty, and -1 complete disbelief. In formal logic  $C(A)$  is restricted to the three values (-1, 0, +1).

Consider the two assertions:

(A)  $3^2 + 4^2 = 5^2$ .

(B) There exist integers x, y, z, such that  $x^2 + y^2 = z^2$ .

If (A) is true, (B) must be also.

These relationships are represented in Modus Ponens as:

$$A \rightarrow B \text{ and } [TV(A) = T; C(A) = 1].$$

$$\text{Therefore, } [TV(B) = T; C(B) = 1].$$

Next, consider Fermat's Last Theorem in the following form.

(A) For any integers  $x, y, z, n$ , ( $n > 2$ ),

$$\text{it is false that } x^n + y^n = z^n.$$

Clearly a single example of integer numbers for which there existed the equality indicated would render (A) false. This mode of reasoning, called Modus Tollens, can be represented as:

$$A \rightarrow B \text{ and } [TV(B) = F; C(B) = -1].$$

$$\text{Therefore } [TV(A) = F; C(A) = -1].$$

It is evident that Ponens and Tollens exhibit unattenuated propagation of T or F in preferred directions (even through chains of implications). Polya pointed out that in normal human reasoning we may make additional logical statements. I call these the Method of Confirmation and the Method of Weakened Support.

Consider assertions like:

(A) The robot hand grasps object1.

(B) Touch sensors in the robot fingers are on.

If assertion (B) is true, that fact tends to confirm (A). The tendency for B to confirm A may be represented logically and Polya did so with non-numeric indicators. In conformity with the notation for Ponens and Tollens, the Method of Confirmation, in the form suggested by Polya, may be stated as:

$$A \rightarrow B \text{ and } [TV(B) = T].$$

Therefore the credibility increment  
propagated to A from B is "more-likely."

For the same two assertions, if the robot should come to believe that (A)

is false, this would tend to make (B) less likely. This kind of reasoning may be described as the Method of Weakened Support. Polya's form is:

$$A \rightarrow B \text{ and } [TV(A) = F].$$

Therefore the credibility decrement  
propagated to B from A is "less likely."

Note that these equations fill the gaps forbidden in formal logic, propagating credibility change where nothing is transmitted in formal logic.

We now introduce a number,  $\Delta$ , to represent the propagated change of belief.  $\Delta$  is a directed quantity (but not a vector) which represents the increment (or decrement) of credibility propagated between the antecedent and consequent of any implication.

$\Delta(A|B)$  is the change of credibility propagated to A from B.

$\Delta(B|A)$  is the change of credibility propagated to B from A.

$$0 \leq \Delta \leq 1.$$

Note also that the syntax of all four equations specifies whether what is transmitted is an increment or decrement. See Table I.

Table I. Syntax of Belief Change

For  $A \rightarrow B$ :

CONFIRMATION	$TV(B) = T$	$\Delta(A B)$ is an increment.
TOLLENS	$TV(B) = F$	$\Delta(A B)$ is a decrement.
PONENS	$TV(A) = T$	$\Delta(B A)$ is an increment.
WEAKENED SUPPORT	$TV(A) = F$	$\Delta(B A)$ is a decrement.

In dealing with multiple support for a given assertion, increments and decrements are added separately by rules defined later, and the resultant sums combined according to:

$$C(A) = I(A) - D(A),$$

where  $I(A)$  is the sum of increments and  $D(A)$  is the sum of decrements.

In order to maintain consistency between the notions of truth of an assertion, and the degree of belief in the truth of that assertion, we introduce a

convention. If credibility becomes positive, the system is to regard the assertion as true with the credibility just established. If credibility becomes negative, the assertion is to be regarded as false with that credibility.

The convention also applies to truth values. The rules are:

For any assertion A,

if  $C(A) > 0$ , set  $TV(A) = T$ .

If  $C(A) < 0$ , set  $TV(A) = F$ .

Conversely, if  $TV(A) = T$ , assert  $C(A) > 0$ .

if  $TV(A) = F$ , assert  $C(A) < 0$ .

Or,  $[TV(A) = T] \equiv [C(A) > 0]$ ,

$[TV(A) = F] \equiv [C(A) < 0]$ .

This equivalence is called the C-TV Consistency rule.

One further definition will permit us to state quantitative equations of Plausible Inference (PI) concisely.

Let  $\Delta l$  be the value of  $\Delta$  for an implication when there is either complete belief or disbelief in either the antecedent or consequent. Then for  $A \rightarrow B$ :

$\Delta l(A|B|+)$  is the value of  $\Delta(A|B)$  when  $C(B) = 1$ .

$\Delta l(A|B|-)$  is the value of  $\Delta(A|B)$  when  $C(B) = -1$ .

$\Delta l(B|A|+)$  is the value of  $\Delta(B|A)$  when  $C(A) = 1$ .

$\Delta l(B|A|-)$  is the value of  $\Delta(B|A)$  when  $C(A) = -1$ .

## 2.1 Quantitative Plausible Inference

Plausible Inference will operate on a dynamic knowledge base. Whenever a state-change takes place, causing either a change in truth value or credibility of an assertion, a new PLANNER-like context layer is pushed. Once such a change or disturbance has taken place, propagation of beliefs occurs along lines of support for logically related assertions within the context layer until a state of equilibrium is reached.

In Plausible Inference logic the values of  $\Delta(A|B)$  or  $\Delta(B|A)$  may be calculated by the following four equations for the inference  $A \rightarrow B$ .

1) Confirmation

If in a new context layer,  $[C(B) > 0]$  is asserted, when previously  $C(B)$  had been zero, then

$$\Delta(A|B) = \Delta 1(A|B|+) \cdot C(B).$$

This means that for our two robot-base assertions A and B, experience may have shown if we are certain that the touch sensors are on, our confidence in the robot grasping something makes  $\Delta 1 = 0.7$ . If, however, those touch sensors are noisy, the robot may have to assign a credibility less than one to "touch sensors are on." This decrease in certainty of touch reduces the credibility increment transmitted to the grasp assertion below 0.7.

2) Graded Tollens

If in a new context layer  $[C(B) < 0]$  is asserted, when previously  $C(B)$  had been zero, then

$$\Delta 1(A|B|-) = 1;$$

$$\Delta(A|B) = \Delta 1(A|B|-) \cdot |C(B)| = |C(B)|.$$

Our credibility decrement to the assertion that the robot is grasping something is equal to the absolute value of the credibility of touch-on being false.

If  $C(B) = -1$ , then  $\Delta(A|B) = 1$ , and  $C(A) = -1$  regardless of other support (unless there is support for  $C(A) = 1$ , which would indicate a contradiction).

3) Graded Ponens

If in a new context layer  $[C(A) > 0]$  is asserted, when previously  $C(A)$  had been zero, then

$$\Delta 1(B|A|+) = 1;$$

$$\Delta(B|A) = \Delta 1(B|A|+) \cdot C(A) = C(A).$$

If other evidence supports the notion that the robot is grasping something, producing a  $C(A) > 0$ , the credibility increment transmitted to touch-on is

equal to  $C(A)$ .

If  $C(A) = 1$ , then  $\Delta(B|A) = 1$ , and  $C(B) = 1$ , regardless of other support (unless there is support for  $C(B) = -1$ , which would indicate a contradiction).

#### 4) Weakened Support

If in a new context layer  $[C(A) < 0]$  is asserted, when previously  $C(A)$  had been zero, then

$$\Delta(B|A) = \Delta(B|A|-) \cdot |C(A)|.$$

Using our robot-base assertions, suppose other evidence supports the notion that the robot has dropped or let go of the object. (It may see an object of similar shape on the floor.) There is a decrement to touch-on transmitted when we are certain that grasp is not true, say 0.8. The decrement transmitted is reduced by the actual credibility of grasp.

As noted above, graded Tollens and Ponens become identical to formal logic when  $C(B) = -1$  or  $C(A) = 1$  respectively, with  $\Delta(A|B|-) = \Delta(B|A|+) = 1$ . To get full formal logic requires, in addition, that  $\Delta(A|B|+) = \Delta(B|A|-) = 0$  for equations 1 and 4, thus blocking transmission of belief in the forbidden directions. It is possible to temporarily reset these particular  $\Delta$ 's to zero for a selected subset of assertions, if there is a need to restrict inferencing to the rules of formal logic for that subset.\*

### 3. Combining Credibilities

We have just considered belief propagation in an implication with a single

---

\*The reader who is familiar with S&B's work can translate our PI notation into S&B's notation by noting that our  $C(A)$  is their  $CF(A,B)$ , our  $\Delta(A|B)$  is their  $MB(A,B)$  or  $MD(A,B)$ , and  $\Delta(A|B|+)$  is  $MB'(A,B)$ . S&B's equations for strength of evidence are identical to our equation 1 for Confirmation.

antecedent and consequent. When dealing with more complex logical expressions with multiple support for antecedents or consequents, S&B's methods for combining beliefs propagated from consequents can be extended to beliefs propagated from antecedents. The following rules include both cases.

### 3.1 Defining Criteria

#### 1) Notation

A is usually an antecedent or a hypothesis.

B is usually a consequent or evidence.

$I(A)$  is defined as the combination of increments of credibility to A [ $I(A) \geq 0$ ].

$D(A)$  is defined as the combination of decrements of credibility to A [ $D(A) \geq 0$ ].

$C(A) = I(A) - D(A)$ .

$E+$  is defined as the set of confirming evidence.

$E-$  is defined as the set of disconfirming evidence.

$A+$  is defined as the set of hypotheses that imply the evidence.

$A-$  is defined as the set of hypotheses that deny the evidence.

$\&$  is defined as logical AND.

$V$  is defined as logical OR.

$\sim$  is defined as logical NOT.

#### 2) Relation between belief and disbelief.

$I(A) = D(\sim A)$ .

In words, the belief in evidence for any proposition A equals the disbelief in evidence against it.

#### 3) Limits

(A)  $\Delta I(B|A|+) = 1$ ; Ponens

$\Delta I(A|B|-) = 1$ ; Tollens

Certainty is transmitted in Ponens and Tollens modes when certain of antecedent or consequent, respectively.

(B)  $0 \leq \Delta(A|B|+) < 1$ ; Confirmation.

$0 \leq \Delta(B|A|-) < 1$ ; Weakened Support.

Certainty is never transmitted in Confirmation or Weakened Support Modes.

(C)  $[C(A) \text{ from } B-] \leq C(A) \leq [C(A) \text{ from } B+]$ .

(D)  $[C(B) \text{ from } A-] \leq C(B) \leq [C(B) \text{ from } A+]$ .

#### 4) Absolute Confirmation or Disconfirmation

(A) If  $\Delta(A|B+) = 1$ , then  $C(A) = 1$ , regardless of  $B-$ .

(B) If  $\Delta(A|B-) = 1$ , then  $C(A) = -1$ , regardless of  $B+$ .

(C) If  $\Delta(A|B-) = \Delta(A|B+) = 1$ , this is contradictory and  $C(A)$  is undefined.

Similar considerations hold for  $\Delta(B|A-)$  and  $\Delta(B|A+)$ .

#### 5) Commutativity

If  $B1$  and  $B2$  indicate an ordered observation of evidence, first  $B1$ , then  $B2$ .

(A)  $\Delta(A|B1 \& B2) = \Delta(A|B2 \& B1)$ .

(B)  $[C(A) \text{ from } B1 \& B2] = [C(A) \text{ from } B2 \& B1]$ .

If  $A1 \vee A2$  indicates an ordered formulation of hypotheses, first  $A1$ , then  $A2$ .

(C)  $\Delta(B|A1 \vee A2) = \Delta(B|A2 \vee A1)$ .

(D)  $[C(B) \text{ from } A1 \vee A2] = [C(B) \text{ from } A2 \vee A1]$ .

#### 6) Missing Information

If  $[TV(B2) = U]$

then  $\Delta(A|B1 \& B2) = \Delta(A|B1)$ .

$[C(A) \text{ from } B1 \& B2] = [C(A) \text{ from } B1]$ .

If  $[TV(A2) = U]$

then  $\Delta(B|A1 \vee A2) = \Delta(B|A1)$ .

$[C(B) \text{ from } A1 \vee A2] = [C(B) \text{ from } A1]$ .

### 3.2 Combining Functions

When combining the increments of credibility of multiply supported assertions, first the increments and decrements of credibility (confirming and



disconfirming) are combined by rules given below, then the difference between confirming and disconfirming beliefs is taken to give the resultant credibility.

In what follows,  $C(A)$  or  $C(B)$  need not be +1 or -1 unless explicitly noted.

### 1) Convergence of Evidence to Support a Given Hypothesis

(COMB1) Assume  $A \rightarrow (B1 \& B2)$

$$\neg A \rightarrow (B3 \& B4).$$

$B1, B2$  are confirming evidence for  $A$ ;  $B3, B4$  are disconfirming evidence.

Suppose  $B1, B2, B3, B4$  are all true.

Then

$$I[A] = \Delta[A|B1] + \Delta[A|B2](1 - \Delta[A|B1]),$$

$$D[A] = \Delta[A|B3] + \Delta[A|B4](1 - \Delta[A|B3]),$$

$$C[A] = I[A] - D[A].$$

As S&B point out, this rule makes the credibility added by the second increment proportional to the remaining disbelief after the first increment is transmitted.

### 2) Convergence of Hypotheses on Given Evidence

(COMB2) Assume  $A1 \rightarrow \neg B$

$$A2 \rightarrow \neg B$$

$$A3 \rightarrow B$$

$$A4 \rightarrow B$$

Let all antecedents be false.

Then

$$I[B] = \Delta[B|A1] + \Delta[B|A2](1 - \Delta[B|A1])$$

$$D[B] = \Delta[B|A3] + \Delta[B|A4](1 - \Delta[B|A3])$$

$$C[B] = I[B] - D[B].$$

### 3) Fanout of Evidence to Conjunctions of Hypotheses

(COMB3a)  $\Delta[(A1 \& A2)|B] = \min(\Delta[A1|B], \Delta[A2|B])$

$$(COMB3b) \Delta[\neg(A1 \& A2)|B] = \max(\Delta[\neg A1|B], \Delta[\neg A2|B]).$$

Belief in a conjunction of hypotheses, from confirming or disconfirming evidence, is limited to the minimum confirmation or maximum disbelief.

4) Fanout of Evidence to Disjunctions of Hypotheses

$$(COMB4a) \Delta[(A1 \vee A2)|B] = \max(\Delta[A1|B], \Delta[A2|B])$$

$$(COMB4b) \Delta[\neg(A1 \vee A2)|B] = \min(\Delta[\neg A1|B], \Delta[\neg A2|B])$$

Belief in disjunctions is complementary to conjunctions.

5) Fanout of an Hypothesis to Conjunctions of Evidence

$$(COMB5a) \Delta[(B1 \& B2)|A] = \max(\Delta[B1|A], \Delta[B2|A])$$

$$(COMB5b) \Delta[\neg(B1 \& B2)|A] = \min(\Delta[\neg B1|A], \Delta[\neg B2|A]).$$

Disbelief in a conjunction of evidence is maximized by loss of support. Belief is minimized by that loss.

6) Fanout of an Hypothesis to Disjunctions of Evidence

$$(COMB6a) \Delta[(B1 \vee B2)|A] = \min(\Delta[B1|A], \Delta[B2|A])$$

$$(COMB6b) \Delta[\neg(B1 \vee B2)|A] = \max(\Delta[\neg B1|A], \Delta[\neg B2|A]).$$

Belief in disjunctions is complementary to conjunctions.

An example may serve to make the application of these rules clearer. Assume the following assertions are in the data base:

(A) The robot hand is attached to object1.

(B1) The robot hand touch sensors are on.

(B2) Finger spread is close to a visually measured dimension D of object1.

(B3) Object1 moves with the hand, or

(B4) Torque or force on the hand saturates when the hand tries to move, and all velocities remain zero.

and  $A \rightarrow (B1 \& B2 \& [B3 \vee B4])$ .

Let  $Z = (B1 \& B2 \& [B3 \vee B4])$ ,

$$Y = \Delta1(A|Z|+).$$

Then  $C(A) = Y \cdot C(Z)$ ,

$$C(Z) = I(Z) - D(Z).$$

Let c.m. stand for confirming measurements, d.m. for disconfirming measurements, m1 for measurement 1, etc. An example of a confirming measurement might be an above-threshold voltage on the touch-sensor line. A disconfirming measurement would be its absence.

$$I(Z) = I(B1 \& B2 \& [B3 \vee B4])$$

$$= \Delta(B1 \& B2 \& [B3 \vee B4] | \text{c.m.})$$

$$= \min(\Delta[B1|m1], \Delta[B2|m2], \Delta[B3 \vee B4 | \text{c.m.}])$$

$$= \min(\Delta[B1|m1], \Delta[B2|m2], \max(\Delta[B3|m3], \Delta[B4|m4]))$$

$D(Z) = \Delta(B1 \& B2 \& [B3 \vee B4] | \text{d.m.})$  is calculated similarly.

#### 4. Revision of Opinions

Polya suggested a number of qualitative rules modeling how we revise already formed opinions in the light of new evidence. These rules were based on his own introspection and examination of the rational behavior of others. Polya stated these rules in the following form:

1) If B is unlikely to be true without A, and B changes to true, there is more than a normal increase in strength of belief in A.

2) If B is supported by many other propositions, B becoming true supports C(A) only weakly.

3) If A is the only justification known for B, and A is false, C(B) is greatly decreased.

4) If A is only one of several justifications for B, A becoming false only slightly weakens belief in B.

##### 4.1 Quantitative Revision Rules

These qualitative rules may be given a quantitative form such as the following

four rules.

(REV1) Assume in context 1,  $(A1 \vee A2 \vee \dots \vee An) \rightarrow B$

and  $TV(A1) = U$ ,

$C(A1) = 0$ ,

$A2, \dots, An$  are false.

By propagation of credibilities to maintain consistency,

$C(B) = -D(B)$ , computed by COMB2 and MISSING INFO;

$TV(B) = F$  by C-TV Consistency.

If  $TV(B) = T$  in context 2,

$C[A1] = I[A1] = \Delta[A1|B] + 0.9|C[B]_{ctx1}|(1 - \Delta[A1|B])$ .

Note: The empirical nature of these rules is illustrated by the factor 0.9, introduced to prevent  $C(A1)$  from reaching +1 in context 2, even if  $C(B) = -1$  in context 1. The factor 0.9 may be regarded as the degree-of-confidence that the system has complete knowledge of the universe of discourse.

In words, REV1 says we increment belief in a hypothesis supported by evidence to the extent that the evidence was previously rendered unlikely by alternative hypotheses.

(REV2) Assume in context 1,  $(A1 \vee A2 \vee \dots \vee An) \rightarrow B$

and  $TV(Ai) = U$ ,  $i = 1, 2, \dots, n$ .

$C(Ai) = 0$ ;  $i = 1, 2, \dots, n$ .

This state of beliefs leaves  $TV(B) = U$ .

If  $TV(B) = T$  in context 2,

then  $\Delta(A1|B) = (1/n)\Delta1(A1|B|+) \cdot C(B)$ .

Note: If in context 2,  $E1 \rightarrow A1$ , and  $TV(E1) = F$ , then  $C(A1)$  is computed by  $I(A1)$  from REV2, and  $D(A1)$  from Weakened Support.

(REV3) Assume in context 1,  $A \rightarrow B$ ,

and no other support, present or potential, exists for B.

Let  $TV(A) = T$ ;  $C[A] > 0$ .

In context 2, let  $TV(A) = F$ . Then

$$C[B] = -D[B] = -[\Delta[B|A] + 0.9 C[A]_{ctx1} \cdot (1 - \Delta[B|A])].$$

The loss of belief in consequence B is proportional to previous confidence in A.

(REV4) Assume in context 1,  $(A_1 \vee A_2 \vee \dots \vee A_n) \rightarrow B$ .

$$TV(A_i) = U; i = 1, \dots, n.$$

Then, if  $TV(A_1) = F$  in context 2, the decrement to B from  $A_1$  is

$$\Delta[B|A_1] = [1/n] \Delta_1[B|A_1] \cdot |C[A_1]|.$$

If  $TV(A_i) = T$ ,  $i > 1$ ,  $C(B)$  is computed by Graded Ponens and REV4.

#### 4.2 A Legal Example of Opinion Revision

Suppose we have the following legal problem, and we are seeking to arrive at a judgment.

A defendant is accused of having blown up the yacht of his girlfriend's father and the prosecution produces a receipt signed by the defendant acknowledging the purchase of some dynamite. Is the defendant guilty?

Assume the following statements are part of the initial contents of the legal knowledge base:

(P1) "D blows up the yacht with dynamite."

(P2) "D acquires dynamite."

(P3)  $(P1 \rightarrow P2)$ .

(P4) "D clears tree stumps."

(P5) "D is a miner."

(P6) "D is a wrecker of buildings."

(P7) "D is a demolition expert."

(P8)  $(P4 \vee P5 \vee P6 \vee P7) \rightarrow P2$ .

##### 4.2.1 Solution to the Legal Problem

We are given in context layer 1,

$$(P1 \vee P4 \vee P5 \vee P6 \vee P7) \rightarrow P2$$

and  $TV(P1) = F; C(P1) = -1; i = 4,5,6,7.$

$$TV(P1) = U; C(P1) = 0; i = 1,2$$

Take, a priori,

$$\Delta(P2|\neg P4) = 0.1$$

$$\Delta(P2|\neg P5) = 0.1$$

$$\Delta(P2|\neg P6) = 0.2$$

$$\Delta(P2|\neg P7) = 0.8.$$

Then, applying COMB2,  $D(P2) = 0.87$

$C(P2) = -0.87$  (highly-unlikely)

$TV(P2) = F$ , by C-TV consistency.

These are the beliefs attained in Context 1.

In context layer 2,

$TV(P2) = T.$

If we assume  $\Delta(P1|P2) = 0.1$  in the absence of other evidence, then applying

REV1,  $C(P1) = 0.1 + (0.9) (0.87) (0.9).$

$C(P1) = 0.80$  (highly-likely).

The unlikelihood of P2 being true without P1 being true makes the truth of P1 highly likely following the discovery that P2 is true.

## 5. Successive Verification

Of particular interest for learning and knowledge acquisition is the case of successive verification of the same hypothesis by similar evidence. (We shall assume that to establish similarity suitable matching criteria are met.) Suppose that the system, by the use of models, has asserted a new implication whose antecedent and consequent may be considered a new hypothesis, A, and evidence for that new assertion, B, respectively. Let the hypothesis A be an assertion to the effect that the performance of some action produces a state whose measurement is

available to the robot, and B is the observation of that measurement. We have the opportunity to assign  $\Delta(A|B)$  automatically so that it conforms to reasonable criteria. If this can be done, it would be an experimentally determined degree-of-belief rather than one supplied by a human expert. We propose the following simple heuristic.

### 5.1 Credibility Growth Through Experience

Assume at the outset that a new implication has been generated by model M, such that  $M \rightarrow (A \rightarrow B)$ , and both A and B are new assertions. We also assume that the truth value of both A and B are measurable by the system in this special case. We introduce a threshold, e, for the system such that

$$\text{for } (-e < \Delta < e), \text{ set } \Delta = 0.$$

Then  $\Delta = e$  is the smallest value of transmissible credibility (a kind of just-noticeable difference). For  $A \rightarrow B$ , set  $\Delta_1(A|B|+) = e$ , before any experience in executing the action described by the hypothesis. We have, as before, that

$$\Delta(A|B) = \Delta_1(A|B|+) \cdot C(B).$$

Assume that  $C(B) = 1$ , during execution of the relevant action.

Let n be the total number of successive trials,

q be the number of trials for which  $[TV(B) = T; TV(A) = T]$ ,

r be the number of trials for which  $[TV(B) = T; TV(A) = F]$ .

Then,

$$(SV1) \Delta_1(A|B|+) = e/n + [(n-1)/n][q/n].$$

$$(SV2) \Delta_1(\neg A|B|+) = e/n + [(n-1)/n][r/n].$$

With these heuristics, for  $n = 1$ , either  $\Delta_1(A|B|+)$  or  $\Delta_1(\neg A|B|+)$  is at most equal to e, the threshold value,

and  $\Delta_1(A|B|+) = \lim_{n \rightarrow \infty} q/n$ , provided the limit exists,

$$\Delta_1(\neg A|B|+) = \lim_{n \rightarrow \infty} r/n, \text{ provided the limit exists.}$$

During the gathering of data,  $q/n + r/n$  does not necessarily equal 1, since

other factors may influence the truth or falsity of A in addition to B's truth; i.e.,  $A \rightarrow B_1 \& B_2 \vee B_3$ , may be the true state of affairs. If this turns out not to be the case, i.e.,  $\lim_{n \rightarrow \infty} q/n + r/n = 1$ , we may presume that a reliable probability has been determined, and  $\lim_{n \rightarrow \infty} q/n = p(A|B)$ .

As S&B show, if  $\Delta 1(A|B|+) = p(A|B)$ , then using their combining rule is consistent with Bayes theorem.

## 5.2 A Robot Learning Example

Suppose the robot has a set of models in its knowledge base like the following:

(MODEL1) For all OBJECTs, there is a threshold pressure, K1, such that if  
(PRESSURE < K1) then

[CAUSES (GRASP OBJECT with PRESSURE)  
(attached OBJECT)]

(MODEL2) For all OBJECTs, there is a threshold pressure, K1, such that if  
(PRESSURE > K1) then

[CAUSES (GRASP OBJECT with PRESSURE)  
(broken OBJECT)]

After attempting to grasp an object, object1, with a normal grasping pressure P1, the robot may detect that it has broken it. The inference mechanism then finds from visually interpreted data that rule MODEL2 matches the caused state of the experience, "broken object1." This enables the inference mechanism to instantiate a rule for objects similar to object1 that places an upper limit on the threshold grasping pressure, K1, in the model rules. It may then use a heuristic to generate a new assertion for grasping similar objects such as:

For all OBJECTs, and all PRESSUREs,  
when [OBJECT is a member of class (object1)] and  
(PRESSURE < P1) then

[CAUSES (GRASP OBJECT with PRESSURE)  
(attached OBJECT)].



By invoking the MODEL heuristics, the robot could learn a reasonable value for the grasp pressure. The robot would then avoid breaking objects that fit the measured description of the broken object. By using rules SV1 and SV2, if the ratios converged, it would simultaneously form a good estimate of the certainty of that knowledge.

## 6. A Model of Human Reasoning

The rules of inference and assumptions introduced here constitute a model of some aspects of human reasoning. The model permits both antecedent and consequent reasoning in either an inexact plausible form or by formal logic when desired. This type of model is designed to imitate belief propagation in bodies of knowledge. It also seems to imitate certain aspects of observed human behavior such as a limited notion of "expertise." For example, one of the differences between a novice and an expert doctor can be represented by supplying the novice model with almost all the relevant facts, but assigning zero or low values to the novice  $\Delta$ s. The novice model will fail to reach many of the conclusions available to the expert model, which has high-confidence values assigned to  $\Delta$ s. This suggests that methods of building up  $\Delta$ s from zero by various reinforcement schemes such as confirmation should be sought after. Learning  $\Delta$ s would be preferable in many circumstances to relying solely on a priori assignment of  $\Delta$ s by an expert. The adoption of PI-type rules of inference does not prevent the use of hierarchical organizations of knowledge to direct the flow of inference. Recent papers employing various rules for assigning confidence in confirmatory evidence have used detailed hierarchical state models of disease for directing inference in diagnosis (Szolovits and Pauker '78; Weiss, Kulikowski, Amarel, and Safir '78).

### 6.1 Fuzzy Interfaces

It is evident that our  $\Delta$  assignments are rough measures of experience, with all its uncertainties. It might be unwise to place much significance in the numbers computed past one or two places. For interfacing with humans, fuzzy set

categories can be defined for ranges of credibility. For example, describe all credibilities  $> 0.8$  as "highly-likely," credibilities between  $-0.1$  and  $0.1$  would be "unknown," etc. Some such minimum threshold of belief rule is probably useful even for internal operations. (If  $-0.1 < C(A) < 0.1$ , then set  $C(A) = 0$ .)

## 6.2 PI in Problem Solving

Plausible Inference permits propagation of belief to take place in directions forbidden to formal logic, thus permitting assertions to become highly likely that are not reachable by formal logic from a given set of assertions. One use of PI may be in combination with formal logic to suggest most promising paths of search for proofs. Polya's thesis in Patterns of Plausible Inference was that even in mathematics, a field where certainties are possible, the mathematician constantly uses plausible reasoning to suggest likely answers before attempting to prove them, and that it is employed in all human reasoning. The problem-solver system being developed at JPL will seek to employ PI in this manner. This problem-solving approach may permit generating command sequences automatically to carry out scientific experiments on board a spacecraft (a kind of automatic programming). PI will also be used in generating robot action sequences designed to carry out high-level commands issued to the JPL robot. Still another application is to RECOGNIZER, a robot problem-solver which learns from experience. RECOGNIZER will be rewritten to incorporate PI.

One final comment is in order. When we quoted Polya at the beginning of this paper, we had a double motive. Of course we are underscoring the manifold developments that may now permit a machine to engage in a form of plausible inference, but in addition, we wish to point up his insight into the essential ingredient—a proper background. The recent development of representations, truth-maintenance, and credibility theory have all contributed to giving the computer that necessary background.

## References

- Doyle, J. '78, Truth Maintenance Systems for Problem Solving, M.I.T., AI-TR-419, January 1978.
- Duda, R. O., Hart, P., Nilsson, N. and Sutherland, G. '77, Semantic Network Representations in Rule-Based Inference Systems, SRI International, AI Center, TN-136, 1977.
- Duda, R. O. et al. '78, Development of the PROSPECTOR Consultation System for Mineral Exploration, Final Report for the period Oct. 1, 1976 to Sept. 30, 1978, SRI International, AI Center, Menlo Park, CA 1978.
- Fikes, R. E. '75, Deductive Retrieval Mechanisms for State Description Models, IJCAI4, Tbilisi, USSR, September 1975, pp. 99-106.
- London, P. E. '78, Dependency Networks as a Representation for Modeling in General Problem Solvers, Ph.D. Thesis, University of Maryland, Computer Science Department, TR-698, September 1978.
- Polya, G. '54, Patterns of Plausible Inference, Princeton University Press, Princeton, NJ, 1954.
- Shortliffe, E. H. and Buchanan, B. G. '75, "A Model of Inexact Reasoning in Medicine," Math. Biosci. Vol. 23, pp. 351-379 (1975).
- Shortliffe, E. H. '76, Computer-Based Medical Consultations: MYCIN, Elsevier, New York, 1976.
- Stallman, R. M. and Sussman, G. J. '77, "Forward Reasoning and Dependency-Directed Backtracking in a System for Computer-Aided Circuit Analysis," Artificial Intelligence, Vol. 9, No. 2, pp. 135-196.
- Szolovits, P. and Pauker, S. G. '78, "Categorical and Probabilistic Reasoning in Medical Diagnosis," Artificial Intelligence, Vol. 11, Nos. 1 and 2, August 1978, pp. 145-172.

(References - Cont'd)

Thompson, A. M. '79, "Semantic Network Truth Maintenance for Deduction and Modeling," JPL, Pasadena, Calif. (Submitted to IJCAI6).

Weiss, S. M., Kulikowski, C. A., Amarel, S., and Safir, A. '78, "A Model-Based Method for Computer-Aided Medical Decision Making," Artificial Intelligence, Vol. 11, Nos. 1 and 2, August 1978, pp. 145-172.